(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 790 304 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
20.08.1997 Bulletin 1997/34

(51) Int. Cl.[6]: **C12N 15/10**, C12N 15/62, C12N 15/81, C12Q 1/68

(21) Application number: **95309340.8**

(22) Date of filing: **21.12.1995**

(84) Designated Contracting States:
AT BE CH DE DK ES FR GB GR IE IT LI LU MC NL PT SE

(71) Applicant: **AMERSHAM INTERNATIONAL plc**
Little Chalfont Buckinghamshire HP7 9NA (GB)

(72) Inventor:
The designation of the inventor has not yet bee n filed

(74) Representative: **Privett, Kathryn Louise**
Stevens, Hewlett & Perkins,
1 Serjeants' Inn, Fleet Street
London EC4Y 1LL (GB)

(54) Improvements relating to assay systems

(57) A method for detecting an interaction between two proteins while eliminating false positive interactions, the method being particularly useful for application in the two-hybrid system and for the screening of libraries of unknown proteins.

EP 0 790 304 A1

Printed by Rank Xerox (UK) Business Services
2.14.12/3.4

Description

The invention is concerned with improvements in methods for detecting protein-protein interactions, more specifically with the reduction or elimination of false positives. The invention relates in particular to improvements to two-hybrid or interaction trap systems for selecting for interacting proteins in living cells. The invention also includes modifications to hybrid constructs and to vectors and yeast strains expressing such constructs.

In the two-hybrid system of Fields and Song(1) described in US 5,283,173 two chimeric genes which encode hybrid proteins are used to test the interaction between a known protein and protein of interest. The first chimeric gene codes for a known protein, often called the bait protein, fused to the DNA-binding domain of a transcriptional activator. The second chimeric gene codes for a protein of interest fused to the transcriptional activation domain. Additionally, the protein of interest may not be known and could be derived for example from a cDNA library. In a suitable host cell such as yeast, if the protein of interest and the bait protein do interact they bring into proximity the DNA-binding and transcriptional activation domains. This proximity is sufficient to cause transcription of a marker gene placed under the control of a promoter containing a binding site for the DNA-binding domain.

Yeast genetic systems have also been used in methods for defining DNA-binding domains of proteins. One such method (8) uses a chimeric protein containing a transcriptional activation domain together with two DNA-binding domains, each capable of binding to a different reporter gene. One of the DNA-binding domains is mutated to analyse its DNA-binding properties.

In the two-hybrid system, once a specific cDNA-encoded protein or known protein of interest has been shown to give rise to activation of the marker gene, it is important to show that this is indeed due to an interaction between the bait protein and the protein of interest and not due to a "false positive" interaction. At least four classes of false positives may occur (2) in two-hybrid systems where a protein of interest, which may be encoded by a cDNA library, is fused to the activation domain:

1. A DNA-binding domain might be cloned into the activator hybrid which would activate transcription independent of the DNA-binding hybrid, by binding:

   a) to the DNA binding site or
   b) to the basal portion of the promoter;

2. Protein in the activator hybrid might bind to the DNA-binding domain of the bait hybrid;
3. Protein in the activator hybrid might bind to the novel junction formed between the protein and an epitope tag or tags encoded by the bait hybrid vector, or to the epitope tag(s) itself;
4. Protein in the bait hybrid might bind to the novel junction formed between the protein of interest and an eptiope tag or tags encoded by the activator hybrid vector.

A number of strategies have been previously described which remove some of the above classes of false positives(2). Three such strategies are as follows:

1. *The use of two marker genes*(2). One of these genes usually expresses a selectable marker (eg *His3*) and the other one a measurable marker activity (eg.*LacZ*) and the reporter gene promoters are usually different. This allows the removal of those proteins in the first class of false positive where a non-specific interaction occurs between the protein and the promoter of the marker gene, since non-specific interaction with both of the marker gene promoters is less likely to occur.
2. *The use of curing to remove the bait hybrid protein*(3). This strategy removes the bait hybrid plasmid, so that only the activator hybrid expressing the protein of interest is present in the cell. If the marker gene expression remains high, then this shows that the activation is due to spurious promoter activation by the protein of interest hybrid protein, rather than via binding through the bait protein hybrid.
3. *The use of an unrelated bait protein*(2). This is to ensure that the protein interacts with the bait protein of interest and no other part of the bait hybrid protein.

Strategies which have previously been applied to the problem of false positives in two-hybrid systems remove only some of the above classes of false positives. In addition, some of these previously used strategies require lengthy and labour intensive screening steps after the initial screen and no strategy has been devised which would be able to remove all of these false positives at the initial screening step. Furthermore, there has so far been no strategy for removing false positives in the screening of two libraries against each other in a search for pairs of interacting proteins. Clearly, there is a need for methods of effectively eliminating false positives in these systems.

The present invention aims to provide [illegible] for detecting interactions between proteins, [illegible] some or all of the problems with known or significant false positives. In particular, the invention aims to [illegible] the present two-hybrid [illegible] each other to make it curing of (removing) provides a method multiple plasmid between a first test protein

and a second test protein, which method comprises:

a) providing a host cell;
b) providing a first chimeric gene that is capable of being expressed in the host cell, the first chimeric gene comprising a DNA sequence that encodes a first hybrid protein, the first hybrid protein comprising:

(i) a first test protein that is to be tested for interaction with a second test protein; and
(ii) a first detection domain;

c) providing a second chimeric gene that is capable of being expressed in the host cell, the second chimeric gene comprising a DNA sequence that encodes a second hybrid protein, the second hybrid protein comprising:

(i) a second test protein that is to be tested for interaction with the first test protein; and
(ii) a second detection domain;

d) providing a control chimeric gene that is capable of being expressed in the host cell, the control chimeric gene comprising a DNA sequence that encodes a control hybrid protein, the control hybrid protein comprising:

(i) the first test protein; and
(ii) a control detection domain;

wherein when the first detection domain and the second detection domain are brought into proximity to one another by an interaction between the first and second hybrid proteins, a first detectable signal is produced, and when the second detection domain and the control detection domain are brought into proximity to one another by an interaction between the second hybrid protein and the control hybrid protein a second detectable signal is produced which is distinguishable from the first detectable signal;
e) introducing the first chimeric gene, the second chimeric gene and the control chimeric gene into the host cell;
f) subjecting the host cell to conditions under which the hybrid proteins are expressed in sufficient quantity for the detectable signals to be produced;
g) determining whether both of the detectable signals are produced. Using such a method, some false positives due to interactions between test proteins and detection domains can be avoided.

The detectable signal may be generated in a variety of different ways. Preferably, th first and second detectable signals are produced by means of activation of first and second reporter genes. A reporter gene can be any gene encoding a detectable polypeptide product, whether it b detectable by means of a positive indicator eg. a colour change, or by selection, eg. on specific growth media.

It is envisaged that the invention may be performed in a variety of different systems. Preferably, the invention is performed as a modified two-hybrid assay in which:

b) (ii) the first detection domain comprises a first DNA-binding domain which recognises a DNA binding site on the first reporter gene;
c) (ii) the second detection domain comprises a transcriptional activation domain; and
d) (ii) the control detection domain comprises a second DNA-binding domain which recognises a DNA binding site on the second reporter gene;

wherein the reporter genes are activated by the transcriptional activation domain when this is in sufficient proximity to the reporter genes, and the DNA binding sites on the reporter genes are different. False positives of type 1 above are thus also eliminated. The DNA binding site is the site generally referred to as the operator in prokaryotic systems, and the upstream activating sequence (UAS) in eukaryotic systems.

Preferably in the modified two-hybrid assay according to the invention, the first and second reporter genes have different promoters as well as different binding sites for the DNA-binding domains. Most preferably, the different promoters have as far as possible no DNA binding motifs in common (all promoters contain a TATA box which binds the TATA factor prior to initiation of transcription by RNA Pol II). This will ensure that all false positives of type 1 are eliminated.

Preferably, also, the first hybrid protein and the control hybrid protein lack any protein binding sites in common other than protein binding sites specific to the first test protein. Protein binding sites are sequences which are capable of taking part in protein-protein interactions. This lack of sequence similarities ensures that false positives falling into classes 2 and 3 above are eliminated.

Preferably, the modified two-hybrid assay according to the invention is performed in yeast, such as *Saccharomyces cerevisiae* or *Saccharomyces pombe*. It is a further preferred feature of the assay that step e), which is the step of introducing the first, second and control chimeric genes into the host cell, is performed by mating haploid yeast cells of opposite mating type, one yeast cell mating type containing the first reporter gene and chimeric genes which encode the first and second hybrid proteins, and an opposite yeast cell mating type containing the second reporter gene and a chimeric gene which encodes the control hybrid protein.

In a second aspect, the invention provides a method as herein described for a modified two-hybrid assay, wherein the first hybrid protein and the control hybrid protein are in the form of a single dual hybrid protein, comprising the first test prot in and the two differ-

ent DNA-binding domains. In such a method, the first and second test proteins may both be derived from cDNA libraries. Th method may thus be particularly useful for screening two cDNA libraries against each other to identify genes coding for proteins that interact with one another, while avoiding false positives of type 1a described above.

In a preferred embodiment according to this second aspect of the invention, and in particular for a library versus library screen, step e), that is the contacting of the dual hybrid protein (first test protein plus two DNA-binding domains) with the second hybrid protein (second test protein plus transcriptional activation domain), is performed by mating haploid yeast cells of opposite mating type, one containing a chimeric gene encoding the dual hybrid protein and at least one of the reporter genes and another containing a chimeric gene encoding the second hybrid protein and at least the other reporter gene. In a particularly preferred embodiment, each of the opposite mating type yeast cells contains both the first and second reporter genes.

The term "two-hybrid assay" as used in connection with the present invention is not intended to be limited to any specific two-hybrid system according to Fields and Song(1). The modified two-hybrid assay according to the invention preferably includes the following general steps:

a) providing a host cell containing a first reporter gene under the control of a first promoter which is activated by a transcriptional activation domain;
b) providing a first chimeric gene that is capable of being expressed in the host cell, the first chimeric gene comprising a DNA sequence that encodes a first hybrid protein, the first hybrid protein comprising:

(i) a first test protein that is to be tested for interaction with a second test protein;
(ii) a DNA-binding domain that recognises a binding site on the first promoter;

c) providing a second chimeric gene capable of being expressed in the host cell, the second chimeric gene comprising a DNA sequence that encodes a second hybrid protein, the second hybrid protein comprising:

(i) a second test protein that is to be tested for interaction with the first test protein; and
(ii) the transcriptional activation domain;

wherein an interaction between the first test protein and the second test protein in the host cell causes the transcriptional activation domain to activate transcription of the reporter gene;
d) introducing the first chimeric gene and the second chimeric gene into the host cell;
e) subjecting th host cell to conditions under which the first hybrid protein and th second hybrid protein are expressed in sufficient quantity for the detectable gene to be activated;
f) determining whether the detectable gene has been expressed;
g) further providing a second reporter gene under the control of a second promoter which is activated by a transcriptional activation domain, said second promoter lacking as far as possible any DNA binding motif in common with the first promoter;
h) further providing a control chimeric gene that is capable of being expressed in the host cell, the control chimeric gene comprising a DNA sequence that encodes a control hybrid protein, the control hybrid protein comprising:

(i) the first test protein;
(ii) a DNA-binding domain that recognises a binding site on the second reporter gene;

(i) introducing the control chimeric gene and the second reporter gene into the host cell;
(j) subjecting the host cell to conditions under which the control hybrid protein and the second hybrid protein are expressed in sufficient quantity for the second reporter gene to be activated; and
(k) determining whether the second reporter gene has been expressed to confirm whether the first and second test proteins interact.

The first and second test proteins as referred to herein, when in the context of the two-hybrid system, may be understood to correspond to the bait protein and the protein of interest respectively. In a library versus library screen, the first test protein is the unknown protein in a hybrid with the DNA-binding domain(s) and the second test protein is the unknown protein in a hybrid with the activation domain.

The chimeric genes encoding the first, second and third hybrid proteins will usually be present in expression vectors such as plasmids. Expression vectors which can easily be transferred between systems are preferred.

Further aspects of the invention include kits containing some or all of the reagents required for performing an assay according to the invention.

A suitable kit according to the invention may c ig prise:

(a) two different bait vectors capable of ent two different hybrid proteins compri each DNA-binding domains and a first te which vector comprising a nucleic acid and an encodes one of the DNA-bindi encoding insertion site for a nucleic ac th first test protein; ressing a dual

or

(b) a single bait vector different DNA- bait hybrid protein

binding domains and a first test protein, the vector comprising nucleic acid sequences which encode the DNA-binding domains and an insertion site for a nucleic acid sequence encoding the first test protein;

and

(c) an activator vector capable of expressing an activator hybrid protein comprising a transcriptional activation domain and a second test protein, the vector comprising a nucleic acid sequence which encodes the activation domain and an insertion site for a nucleic acid sequence encoding the second test proteins;

and

(d) two reporter genes capable of producing different detectable signals, each reporter gene having a DNA binding site for one of the DNA-binding domains encoded by the bait vector or vectors.

Alternatively, the activator or bait vectors, or both, may already contain sequences encoding the test proteins, eg. from a cDNA library.

The reporter genes in the kit according to the invention may be contained in plasmids. Preferably however they are integrated into yeast cells of opposite mating type.

The term "test protein" as used herein is considered to include peptides, polypeptides and related molecules including for example antibody binding fragments. One or both of the test proteins may consist of a fragment of a particular protein of interest.

Although the preferred host cells for use in the assay according to the invention are yeast cells, the invention may also be applied to two-hybrid screens in mammalian cells, insect cells, plant cells or any intact eukaryotic or prokaryotic organism. In particular with use of yeast cells, the assay method will be capable of automation. It is envisaged that the removal of false positive interactions and the speed and ease of mating between two haploid yeast cells will make it feasible to perform analysis of protein-protein interactions with higher throughput. This makes automation an attractive facility.

The use of two haploid yeast strains is a preferred but not essential part of the invention. In an alternative embodiment, a whole screening system may be set up in a single haploid or diploid strain which contains at least two reporter genes and has been transformed with the appropriate chimeric genes.

Other types of two-hybrid system can be envisaged which do not depend for their end point on transcriptional activation. The principle of the present invention as applied to two-hybrid systems, that different hybrid proteins comprising the same bait protein are screened so that activation of two measurable end points results, is also applicable to these types of two-hybrid assays.

When transcriptional activation of reporter genes is employed in the method according to the invention to provide th necessary detectabl signals, at least two reporter gen s with different DNA binding sites are required. Further reporters may also be used, which can be of any suitable type (measurable or selectable). Overall, the reporter promoters may share sequences in common, provided of course that no single DNA binding site is present in all of the reporter gene promoters which are employed.

Suitable reporter genes will be known to those skilled in the art. Reporter systems which may be used in the invention include but are not limited to anxotropic markers such as Histidine, β-galactosidase, luciferase, green fluorescent protein and Ubiquitin (9). An advantage of the ubiquitin system is that it allows monitoring of protein-protein interactions as a function of time. Also, the interactions can occur in the cytoplasm as opposed to the nucleus (in the conventional two-hybrid approach, the interaction must occur in the nucleus).

Suitable detection domains for the hybrid proteins in the method according to the invention have been described so far only with reference to two-hybrid assays. Other suitable detection domains envisaged include fluorescent moieties or moieties capable of becoming fluorescent. Such detection domains may work in the invention in the following way. The first and second detection domains, when in sufficiently close proximity, cause the emission of light at a particular wavelength. The second detection domain and the control detection domain, when in sufficiently close proximity to one another, cause light of a different wavelength to be emitted. The first and control detection domains may thus each be a different fluorescent moiety for which the wavelength is altered when the moiety is in close proximity with the second detection domain. Alternatively, the second detection domain is a fluorescent moiety for which the wavelength of the light emitted is altered to a different extent by proximity to the first and second detection domains respectively. In this type of system the first and second detectable signals are light emitted at different wavelengths.

The term DNA binding motif is used herein to describe a DNA binding site to which a protein, such as a DNA-binding domain in the method according to the invention, is capable of binding. Suitable reporter genes, promoters, DNA-binding and activation domains etc. for use in the invention will be known to those skilled in the art.

Previously, three DNA-binding domains have been commonly used in variants of the two-hybrid or interaction trap system. These include the Gal4 DNA-binding domain (1), the LexA DNA-binding domain (4), and the serum response factor DNA-binding domain (5). These DNA-binding domains have never previously been used together in a dual bait two - hybrid system nor are there any examples of two DNA-binding domains being used together in a two-hybrid system for removal of false positives.

In detail the invention may be used practically in the following way, which exemplifies the principles of the invention described above:

In the first part of the false positive elimination system, the principles of present state of the art two-hybrid systems are applied. This consists of the use of two marker genes, a selectabl marker (eg.*His3*) and a measurable marker (eg. *LacZ*). These marker genes are placed in a haploid yeast strain under the control of promoters containing a specific upstream activating sequence (UAS) or DNA binding site (eg. the Lex DNA binding sequence (4)). This sequence is placed adjacent to a basal promoter sequence which lacks a UAS (eg. basal sequences from the CYC1, GAL1 or HIS promoters). These marker constructs will preferably be integrated into the yeast genome, but this is not a necessary part of the invention. The strain is transformed with a chimeric bait construct so that it expresses a bait hybrid protein comprising the bait protein domain fused to a DNA-binding domain (eg. the LexA DNA-binding sequence). The strain is also transformed with a further library of constructs which contain a set of chimeric activator genes, expressing a cDNA library fused to the activator sequence (eg. the Gal4 activator sequence). In this strain those colonies which express both the measurable and selectable marker genes will be expressed. The use of a single strain, with two marker genes for identification of two-hybrid interaction proteins expressed from single activator and bait constructs is representataive of the state of the art prior to the present invention.

For the present invention a second haploid yeast strain of opposite mating type is a preferred component. This strain contains a marker gene integrated into its genome. This marker gene could be either a selectable or measurable marker gene (or two marker genes representing one of each type could also be used). In the preferred example a measurable marker gene is used, such as the green fluorescent protein GFP or a luciferase-encoding gene. The marker gene is the second strain is under the control of a different basal promoter sequence to the marker genes in the first strain (eg. Gal1 instead of CYC1). This promoter also contains a different UAS which is a specific DNA binding site (such as SRE instead of Lex). This means that there is as far as possible no similarity in sequence between the promoters. It is a preferred feature of this invention that while any two promoters may share sequence similarity, no single sequence element is present in each of the promoters in front of the marker genes. A second bait hybrid protein will be expressed in the second strain which contains a different DNA-binding domain which is specifically able to bind to the DNA binding site in front of the third marker gene (ie. SRF which binds to the SRE promoter DNA sequence). Any epitope tags or intervening coding sequences between the DNA-binding domain and the bait protein will preferably be different to the first bait hybrid protein so that apart from the bait protein there are as far as possible no ther sequence similarities between the two bait proteins. The second yeast strain may be mated to the first strain to produce a diploid strain. If the protein of interest participates in a specific interaction with the bait protein, then the diploid strain will express all three reporter genes. Since there are no protein sequences in the bait hybrid (apart from the specific bait protein sequence of interest) or DNA sequences in common between all three promoters, then only a true positive can give rise to expression of all three reporter genes. This strategy removes each of the classes of false positives that arise from interactions between molecules other than the protein of interest and the bait protein.

As well as screening for false positives in cDNA library screens in the two-hybrid system this approach will have particular utility when screening random peptide libraries (7) and antibody binding site libraries. Such libraries contain peptides of random sequence composition which may be screened in the two-hybrid system in order to identify those which bind specifically to the bait protein of interest. Since they are random, there is an equal chance of a peptide occurring which binds to the bait protein domain of interest as there is of a peptide occurring which binds to the DNA-binding domain. The present invention will remove such false positives efficiently.

A further application of the invention to the two-hybrid system is for screening two libraries against each other in order to identify pairs of unknown interacting proteins. This approach may have a major application for identification of unknown gene function resulting from the human genome sequencing and mapping projects. Here, the removal of false positives is critical due to the large number of anticipated protein interactions which will be identified. Since in this approach the bait protein is unknown and is replaced by a second cDNA library, it is not possible to prepare matching strains transformed with bait chimeric constructs encoding pairs of bait hybrid proteins. Therefore, a modification to the above invention is described below which is based on the same novel principle described above.

For screening two libraries, two yeast strains are required which are similar to those described above. They encode a total of at least three marker genes, at least two of which have promoters with entirely different sequences placed in front of the marker genes. A first cDNA library is fused to the activator domain and the resulting set of chimeric genes are transformed into the first haploid strain. No activation of any marker gene should occur in this strain at this stage since no bait is present. A second cDNA library is then fused to tw DNA-binding domains joined together in the sam tein to prepare a dual hybrid bait protein. This transformed into the second bait strain. As number of clones in the library m low sequences encoding proteins w strain tional activation domains of the use rise to reporter activation expression encoding clones of th activation of the should then of a sel

can be selected against (such as ura3 which can be selected against using 5-FOA). The two haploid yeast strains can then be mated and resulting diploid strains which express each of the marker proteins selected for and analysed. This will give rise to a bank of yeast colonies expressing interacting pairs of proteins which are substantially free from false positives. False positives which may occur in the library vs library screening strategy are those activator library proteins that bind to the DNA-binding portion of the chimeric bait protein, or possibly bait proteins that bind to the activator domain of the chimeric activator protein. Further analysis using the preparation of single bait hybrid proteins from the known proteins as described above, will largely exclude this class of false positives.

The elimination of false positives is essential for the analysis of interacting proteins by library versus library screening in the two-hybrid system. Efficient removal of false positives will permit high throughput automation of the screening process.

False Positives

The present invention is directed in particular to the screening of DNA libraries against known proteins, or against DNA libraries, to detect interacting proteins. The false positives which may arise when using a known bait protein to screen a library have been discussed above, numbered 1 to 4. The additional false positives which may be encountered when one library is screened against another are:

5. A transcriptional activator might be cloned into the DNA-binding hybrid which would activate transcription independently of the activator hybrid;
6. Protein encoded by the bait hybrid might bind to the activation domain of the activator hybrid.

Also covered by the invention are schemes for reducing or eliminating false positives, which will be described below. These may apply to library versus library screening only, or single library screening also. All of these schemes are based on the two-hybrid system.

I. Removal of false positives from libraries prior to mating

(i) Removal of DNA-binding domains and transcriptional activation domains from library clones.

This scheme removes false positives from groups 1(a), 1(b) and 5 in the list above. Each strain of opposite mating type contains a reporter such as ura3 that can be selected against. The library hybrids can be intr - duced into the strains and any clones which activat the ura3 reporter can be selected against using 5'FOA. The remaining viabl yeast clones can be mated to look for true positives. Strains of opposite mating type containing library hybrids are shown in Figure 1a.

(ii) Removal of DNA-binding domains and transcriptional activation domains from library clones together with any clones which bind to the DNA-binding domain of the bait hybrid vector and any clones which bind to the activation domain of the activator vector.

This scheme involves an empty bait (i.e. DNA-binding hybrid minus the library protein) and an empty activator (i.e. activation hybrid minus the library protein). This is a further development of (i) above which removes false positives in classes 1(a), 1(b), 2, 5 and 6. It could be used in both a single library screen and a library versus library screen. Strains of opposite mating type for this scheme applied to library versus library screening are shown in Figure 1b.

II. An alternative scheme for removing DNA-binding domains and transcriptional activation domains from library clones post mating

This scheme involves the use of an inducible yeast promoter driving the activator hybrid. A key point is that the two reporters have the same basal promoter such as Gal 1 but different specific DNA binding sites. Note that the bait has to contain a single DNA-binding domain for this scheme because it relies on a second reporter to remove false positives as described below.

A true interaction will result in expression of reporter 1 but not reporter 2 under inducing conditions. In non inducing conditions, both reporters would be switched off.

A false positive is present if the first reporter is expressed when cells are grown under non inducing conditions. In this situation, the activator hybrid is "off" or expressed at very low levels. It is possible to screen out activation domains inadvertently cloned as part of a library into the bait hybrid vector by plating the cells with negative selection under non inducing conditions.

In the alternate case there may be general DNA-binding clones in the activator hybrid library which bind to the basal reigon of the promoter e.g. TATA binding proteins. A false positive would result in expression of reporter 2, under inducing conditions.

If the second reporter can be selected against e.g. ura3, then both sets of false positives described above could be removed from the libraries. False positives in classes 1(a), 1(b) and 5 are removed.

This scheme is illustrated in Figure 2.

EXAMPLES

Applicati n of this Inventi n:

The invention will allow the removal of two principle classes of false positives which are not fully removed by current two-hybrid systems without time-consuming

methods which involve plasmid curing and plasmid rescue.

1. Those proteins which bind to a specific element in the promoter. In particular, those proteins which interact with the UAS sequence (eg. Lex or SRE).
2. Those proteins which interact with the DNA-binding domain of the bait protein or other sequences present in the bait protein which are not specifically present in the bait protein domain of interest.

The use of two DNA-binding domains in dual bait proteins allows the use of two or more independent reporter genes which do not have promoter sequences in common. The Examples which follow show that this principle can be applied to removing both classes of false positive.

Example 1

Removal of a false positive arising from a protein which interacts with the reporter gene promoter sequence:

The SRF (serum response factor) binds to the serum response element. Using the SRE-SRF interaction as a DNA-binding domain-promoter UAS combination, the SRE sequence is placed in front of a Gal1 basal promoter and a reporter gene such as GFP. The SRF DNA-binding domain is placed in a bait chimeric vector so that it may be expressed as a fusion protein with any bait protein of interest. The bait protein used in this example is the Pho80 protein from yeast. This protein alone will not give rise to transcriptional activation of the GFP reporter. An activator chimeric construct or library is then introduced into the strain. Proteins which interact with the bait protein domain of interest will then give activation of the GFP reporter. However, if a mammalian cDNA library is used, then SRE binding proteins such as SRF will be present in the library. Such proteins will activate transcription of the GFP reporter gene since they will be fused to a transcriptional activation domain such as that from VP16 or Gal4. To illustrate this, a deliberately constructed chimeric construct expressing SRF fused to the VP16 or Gal4 activation domain is shown to give activation of the GFP reporter through the SRF-SRE interaction. However, when a dual assay is used using a LexA DNA-binding domain in a second bait, and LexA UAS -His3 and LacZ reporters, this does not give rise to activation of the other two reporters. Therefore a "false positive" of this type would be removed at the first screening stage of a two-hybrid assay. In this example, SRF and LexA are the first detection domain and the control detection domain respectively in the method according to the invention.

Example 2

Removal of false positives arising from interactions between the protein of interest and backbone sequences in the bait protein ther than the bait protein domain of interest.

Pho80 expressing bait containing the SRF DNA-binding domain is prepared as above. This protein will bind to the SRE element in the promoter in front of the GFP reporter as above. If a protein is expressed from the activator which binds to SRF, then this protein binds to the SRF DNA-binding domain, rather than the Pho80 protein of interest in the bait, giving rise to false positive activation of expression of the GFP reporter. SAP1 is an SRF binding protein which gives rise to such a false positive result. To illustrate this, SAP1 is cloned into a two-hybrid chimeric activator construct where it is expressed as a fusion with the Gal4 activation domain. In a single bait two-hybrid experiment this construct gives rise to "aberrant" GFP expression. By using the double bait approach with the dual Lex reporters, the Sap1 is not capable to give rise to activation of the Lex-regulated His3 and LacZ reporter genes. Once again therefore a "false positive" of this type would be eliminated at the first screen for true two-hybrid interactions.

REFERENCES

1. Fields,S. and Song,O.-K.(1989) *Nature*340: 245-246
2. Bartel,P., Chien,C.-T., Sternglanz,R. and Fields,S.(1993) *Biotechniques*14: 920-924
3. Harper,J.W., Adami,G.R., Wei,N., Keyomarsi,K. and Elledge,S.J., (1993) *Cell*75: 805-816
4. Gyuris,J., Golemis,E., Chertkov,H., and Brent,R. (1993) *Cell*75: 791-803
5. Dalton,S. and Treisman,R. (1992) *Cell*68: 597-612.
6. Bendixen,C., Gangloff,S. and Rothstein,R. (1994) *Nucl.Acids Res.* 22: 1778-1779.
7. Yang,M., Wu,Z. and Fields,S. (1995) *Nucl.Acids Res.* 23: 1152-1156.
8. Wilson,T.E., Padgett,K.A., Johnson,M. and Milbrandt,J. (1993) *Proc.Natl.Acad.Sci.USA* 90: 9186-9190.
9. Johnsson,N. and Varshavsky,A. (1994) *Proc. Natl. Acad.Sci.USA.* 91:10340-10344.

Claims

1. A method for detecting an interaction between a first test protein and a second test protein, which method comprises:

   a) providing a host cell;
   b) providing a first chimeric gene that is capable of being expressed in the host cell, the first chimeric gene comprising a DNA sequence

that encodes a first hybrid protein, the first hybrid protein comprising:

(i) a first test protein that is to be tested for interaction with a second test protein; and
(ii) a first detection domain;

c) providing a second chimeric gene that is capable of being expressed in the host cell, the second chimeric gene comprising a DNA sequence that encodes a second hybrid protein, the second hybrid protein comprising:

(i) a second test protein that is to be tested for interaction with the first test protein; and
(ii) a second detection domain;

d) providing a control chimeric gene that is capable of being expressed in the host cell, the control chimeric gene comprising a DNA sequence that encodes a control hybrid protein, the control hybrid protein comprising:

(i) the first test protein; and
(ii) a control detection domain;

wherein when the first detection domain and the second detection domain are brought into proximity to one another by an interaction between the first and second hybrid proteins, a first detectable signal is produced, and when the second detection domain and the control detection domain are brought into proximity to one another by an interaction between the second hybrid protein and the control hybrid protein a second detectable signal is produced which is distinguishable from the first detectable signal;
e) introducing the first chimeric gene, the second chimeric gene and the control chimeric gene into the host cell;
f) subjecting the host cell to conditions under which the hybrid proteins are expressed in sufficient quantity for the detectable signals to be produced;
g) determining whether both of the detectable signals are produced.

2. A method as claimed in claim 1, wherein there is further provided in the host cell, a first reporter gene which is activated to produce the first detectable signal when the first and second hybrid proteins interact, and a second reporter gene which is activated to produce the second detectable signal when the first and control hybrid proteins interact.

3. A method as claimed in claim 2, which is a modified two-hybrid assay in which:

b) (ii) the first detection domain comprises a first DNA-binding domain which recognises a DNA binding site on th first reporter gene;
c) (ii) the second detection domain comprises a transcriptional activation domain; and
d) (ii) the control detection domain comprises a second DNA-binding domain which recognises a DNA binding site on the second reporter gene;

wherein the reporter genes are activated by the transcriptional activation domain when this is in sufficient proximity to the reporter genes, and the DNA binding sites on the reporter genes are different.

4. A method as claimed in claim 3, wherein the first and second reporter genes have different promoters.

5. A method as claimed in claim 3 or claim 4, wherein the first hybrid protein and the control hybrid protein have no or substantially no protein binding sites in common other than protein binding sites specific to the first test protein.

6. A method as claimed in any one of claims 1 to 5, for screening proteins encoded by a DNA library, wherein the second test protein is encoded by a DNA sequence from a cDNA library.

7. A method as claimed in claim 6, wherein the host cell is yeast.

8. A method as claimed in any one of claims 2 to 5, wherein step e) is performed by mating haploid yeast cells of opposite mating type, one containing the first reporter gene and chimeric genes encoding the first and second hybrid proteins and another containing the second reporter gene and a chimeric gene encoding the control hybrid protein.

9. A method as claimed in claim 3 or claim 4, wherein the first hybrid protein and the control hybrid protein are in the form of a single dual hybrid protein, comprising the first test protein and the two different DNA-binding domains.

10. A method as claimed in claim 9, wherein th and second test proteins are derived f libraries. of

11. A method as claimed in claim taining step e) is performed by m brid pro- opposite mating typ gen ncoding th and second a chimer tein. first

12. A kit for detecting an interaction between pairs of test proteins, which kit comprises:

(a) two different bait vectors capable of expressing two different hybrid proteins comprising different DNA-binding domains and a first test protein, each vector comprising a nucleic acid sequence which encodes one of the DNA-binding domains and an insertion site for a nucleic acid sequence encoding the first test protein;

or

(b) a single bait vector capable of expressing a dual bait hybrid protein comprising two different DNA-binding domains and a first test protein, the vector comprising nucleic acid sequences which encode the DNA-binding domains and an insertion site for a nucleic acid sequence encoding the first test protein;

and

(c) an activator vector capable of expressing an activator hybrid protein comprising a transcriptional activation domain and a second test protein, the vector comprising a nucleic acid sequence which encodes the activation domain and an insertion site for a nucleic acid sequence encoding the second test proteins;

and

(d) two reporter genes capable of producing different detectable signals, each reporter gene having a DNA binding site for one of the DNA-binding domains encoded by the bait vector or vectors.

Figure 1a

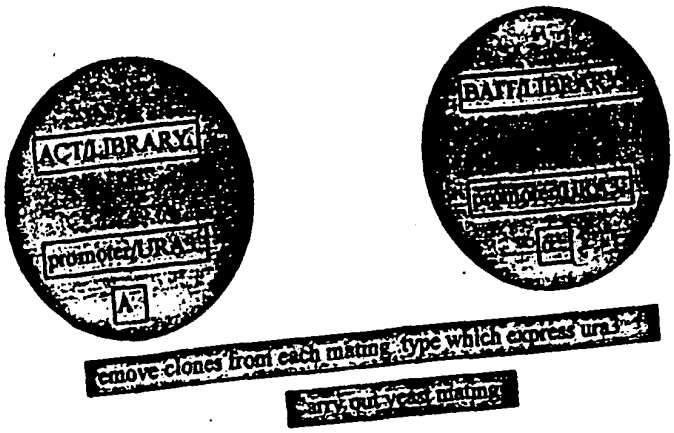

Figure 1b

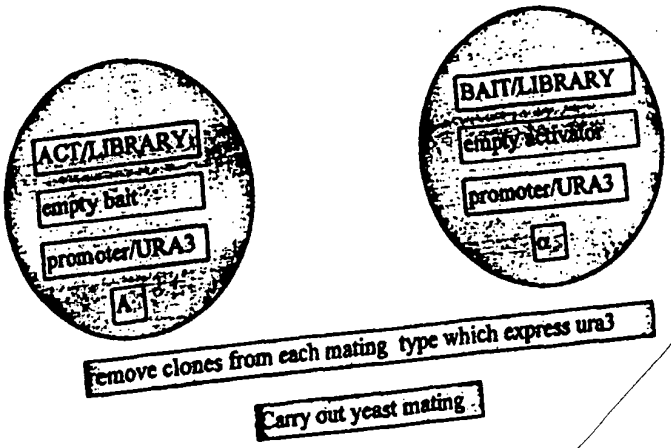

Figure 2


inducible promoter/activator hybrid(s)
LexA/bait hybrid
lexA/Gal basal p/reporter 1
[SRE(s)/Gal basal p/reporter]

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 95 30 9340

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
D,A	BIOTECHNIQUES 14 (6). 1993. 920-922, 924. CODEN: BTNQDO ISSN: 0736-6205, XP002004036 BARTEL P ET AL: "ELIMINATION OF FALSE POSITIVES THAT ARISE IN USING THE TWO - HYBRID SYSTEM." * the whole document * ---	1-12	C12N15/10 C12N15/62 C12N15/81 C12Q1/68
A	CURRENT OPINION IN BIOTECHNOLOGY 6 (1). 1995. 59-64. ISSN: 0958-1669, XP000571534 LUBAN J ET AL: "The yeast two - hybrid system for studying protein-protein interactions." * the whole document * ---	1-12	
A	WO-A-95 20652 (MEDIGENE GMBH ;ALTMANN HERBERT (DE); WENDLER WOLFGANG (DE)) 3 August 1995 * the whole document * ---	1-12	
A	WO-A-94 10300 (GEN HOSPITAL CORP) 11 May 1994 * the whole document * ---	1-12	
D,A	NUCLEIC ACIDS RESEARCH, vol. 22, no. 9, 1994, IRL PRESS LIMITED,OXFORD,ENGLAND, pages 1778-1779, XP002004038 C. BENDIXEN ET AL.: "A yeast mating selection scheme for detection of protein-protein interaction?" * the whole document * ---	1-12	

-/--

TECHNICAL FIELDS SEARCHED (Int.Cl.6)
C12N
C12Q

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	28 May 1996	Hornig, H

CATEGORY OF CITED DOCUMENTS
X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 95 30 9340

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
D,A	PROC. NATL.ACAD SCI., vol. 90, October 1993, NATL. ACAD SCI.,WASHINGTON,DC,US;, pages 9186-90190, XP002004039 T.E. WILSON ET AL.: "A genetic method for defining DNA-binding domains: Application to the nuclear receptor NFGFI-B" * the whole document * ---	1-12	
D,A	CELL, vol. 68, February 1992, CELL PRESS,CAMBRIDGE,MA,US;, pages 597-612, XP002004040 S. DALTON AND. R. TREISMAN: "Characterization of SAP-1, a protein recruited by serum response factor to the c-fos serum response element" * the whole document * -----	1-12	
			TECHNICAL FIELDS SEARCHED (Int.Cl.6)

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
THE HAGUE	28 May 1996	Hornig, H

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)